Available online at www.sciencedirect.com
ScienceDirect
Procedia Computer Science 35 (2014) 986 – 993
Procedia Computer Science

18th International Conference on Knowledge-Based and Intelligent Information & Engineering Systems - KES2014

Estimating collaborative attitudes based on non-verbal features in collaborative learning interaction

Yuki Hayashi[a,*], Haruka Morita[b], Yukiko I. Nakano[b]

[a]*Osaka Prefecture University, 1-1 Gakuen-cho, Naka-ku, Sakai-shi, Osaka 599-8531, Japan*
[b]*Seikei University, 3-3-1 Kichijoji-kitamachi, Musashino-shi, Tokyo 180-8633, Japan*

Abstract

To understand collaborative learning interaction, it is important to analyze not only argument processes based on verbal information but also non-verbal interaction. In order to analyze learning situations in collaborative learning, our previous work proposed an estimation method for learning attitudes based on participants' non-verbal features. Because the method used limited features, this research enhances the method of the participants' collaborative attitudes by analyzing non-verbal features in detail. The model also considers participants' knowledge of their learning subject in the analysis. The estimation model detects three levels of the participants' collaborative attitudes based on multinomial logistic regression analysis. The results of the analysis show that the speech interval feature, in particular, affects the participants' collaborative attitudes. In addition, the results indicate that speakers with knowledge of the learning subject receive more attention from participants with insufficient knowledge. The results of the model evaluation find that the f-measure for classifying the participants' collaborative attitudes is 0.569; for participants with knowledge, the f-measure is 0.647.

Peer-review under responsibility of KES International.

Keywords: Collaborative learning; collaborative attitudes; nonverbal features; multinominal logistic regression.

1. Introduction

Collaborative learning is a learning style where multiple participants study to acquire knowledge of their learning subjects[1]. The fact that participants obtain a wide variety of educational benefits is supported by many theories,

* Corresponding author. Tel.: +81-72-254-9672.
E-mail address: hayashi@kis.osakafu-u.ac.jp

Peer-review under responsibility of KES International.
doi:10.1016/j.procs.2014.08.184

mainly through the collaborative interaction process[2]. However, there are also studies that indicate that collaborative learning does not always bring desirable effects to all participants[3] (e.g., participants might arrive at irrational conclusion if only following the opinion of an egotistic participant). Therefore, technology that can evaluate collaborative learning situations, such as one that determines whether participants truly study collaboratively, is important for judging the outcome of collaborative learning.

In collaborative learning performed in face-to-face environments, the participants not only progress in their learning activity by exchanging utterances, but also by transmitting non-verbal information, such as looking at other participants. Studies have demonstrated that interaction is important not only for daily conversation, but also for collaborative learning to maintain smooth communication[4, 5]. However, to the best of our knowledge, few existing studies focus on analyzing collaborative learning from the perspective of non-verbal information. Whereas several studies have been conducted to analyze interaction in terms of non-verbal features in the field of human-computer interaction[6, 7], in order to analyze collaborative learning, such learning aspects as the difference between the participants' level of knowledge in their learning subjects should be considered.

In order to clarify the role of non-verbal features in collaborative learning interaction, this study analyzes such non-verbal interaction among participants. Discussions with others through collaborative means are crucial to attaining successful learning in collaborative learning. Therefore, we focus in particular on analyzing and estimating whether the participant attempted to advance the discussion collaboratively (*collaborative attitudes*). First, to analyze the non-verbal features that relate to the participants' collaborative attitudes, we conduct an experiment to score each of the participant's collaborative attitudes using the multimodal corpus in collaborative learning collected by our previous research[8]. Then, we verify the effects of non-verbal features and propose an estimation model of the collaborative attitudes compared to the scores based on multinomial logistic regression analysis.

2. Previous work

In order to analyze collaborative learning in terms of non-verbal information, we developed a collaborative learning environment to collect non-verbal information using multimodal measurement devices[8]. Figure 1 represents our collaborative learning environment. Figure 2 shows a layout of the participants and the multimodal measurement devices. The participants are arranged in a triangle formation around a square table. Each participant wears eye tracking glasses and a microphone, and he/she takes notes using a digital pen device. This information is stored into plural computers with timestamp. In addition, infrared (IR) markers with unique IDs are placed on each participant and set of notes to detect the gaze targets based on the eye tracking glasses.

In the learning environment, we conducted experiments to collect multimodal information during collaborative learning. For the experiments, 30 participants contributed in collaborative learning groups. Each learning group consisted of three participants, and ten groups were created. Each group was arranged such that they contained participants who were familiar (two participants: *A* and *B*) and unfamiliar (one participant: *C*) with the learning subjects. In the experiment, the group members were asked to study together for two sessions. We set two types of exercises: an exercise in which the participants took notes frequently to derive a unique answer (type 1); and an exercise in which participants mainly discussed and shared knowledge with other participants (type 2). The discussion time for all the exercises lasted approximately 10 min. We observed the learning progress and stopped the discussion when the conversation quieted down. The collected data was annotated to correct gaze targets and writing action intervals, and used to construct an interaction corpus that includes plenary annotated gaze targets, speech intervals, and writing actions of the participants in the collaborative learning.

Based on the multimodal corpus, we propose a visualization system to briefly analyze interaction sequences during collaborative learning[9]. In the system, and in order to represent the participants who develop a collaborative attitude toward others, each participant's collaborative attitude is estimated as increasing at a rate proportional to the amount of speech time and the amount of time that the other participants gaze at the speaker. The experimental results show that there is a significantly strong correlation between the collaborative attitudes calculated by the system according to non-verbal features, and the subjective judgment of human subjects. However, the model does not consider other non-verbal features, such as the participants' own gazing features. In addition, the effect of the difference in the group's constitution on its members' collaborative attitudes, such as whether a participant already knows their learning subject, remains unsettled. This paper is intended as an investigation of the relationship

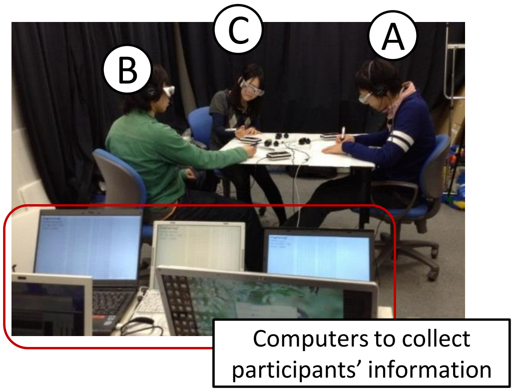

Fig. 1. Snapshot of collaborative learning environment

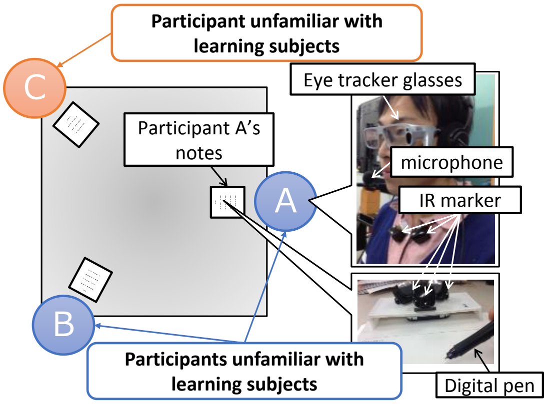

Fig. 2. Layout of participants and multimodal measurement devices

between non-verbal features and collaborative attitudes of participants in more detail based on the multimodal corpus so as to enhance the estimation model.

3. Scoring participants' collaborative attitudes

3.1. Experimental outline

In order to analyze the participants' non-verbal features in the multimodal corpus and their collaborative attitude, we conducted experiments to score collaborative attitudes. In these experiments, we focused on the discussion-type sessions (type 2, ten sessions) as the evaluation target data because this type of exercise is much more conducive to interactions while discussing and sharing knowledge, such as utterances and eye movements, compared to the type 1 exercise, in which participants tend to spend time writing. The participants' discussion topics and behavior might vary with time. Therefore, we divided each video of the learning sessions into two; one video contains 5 min of the first half of the original video, and the other contains 5 min of the last half. The videos were segmented for only 5 min of the first half if the session lasted less than 10 min. As a result, we prepared 17 session videos for evaluation.

Ten scorers participated in the experiment. Four to six scorers were assigned per session video. First, the scorers adequately observed the three participants in the video because we allowed them to watch the video repeatedly. Then, the scorers assigned a point (one for worst to ten for best) to each participant in the video according to whether the participant attempted to advance the discussion through a collaborative attitude. Here, the scorers had to choose a different point for each participant in the same video. We assigned the videos to the scorers so that no scorer evaluated the same group session. In addition, and in order to avoid the order effect, the session videos were observed in random order.

3.2. Results

Table 1 lists the average scores of each participant's collaborative attitude. Here, the notations *fh* and *lh* that appear after the session number represent the first half and the last half of the videos, respectively. The number in parentheses indicates the number of scorers that observed the video. In Table 1, *A* and *B* represent the participants

Table 1. Average scores for collaborative attitude

Session	*A*	*B*	*C*	Session	*A*	*B*	*C*
1-*fh* (6)	7.50	5.33	4.67	7-*fh* (4)	6.25	5.75	5.50
2-*fh* (4)	6.00	5.50	3.25	7-*lh* (4)	4.00	3.25	6.75
2-*lh* (4)	5.75	4.75	2.25	8-*fh* (4)	1.75	7.00	3.50
3-*fh* (4)	6.50	6.25	6.25	8-*lh* (4)	4.00	7.00	4.50
3-*lh* (4)	5.50	2.00	3.25	9-*fh* (4)	5.50	6.75	4.25
4-*fh* (6)	7.50	7.25	4.67	9-*lh* (4)	3.00	2.50	5.50
5-*fh* (4)	5.50	4.75	4.75	10-*fh* (4)	5.75	5.25	4.75
5-*lh* (4)	7.25	4.75	5.50	10-*lh* (4)	6.00	5.00	4.25
6-*fh* (6)	5.33	6.50	4.17				

with knowledge of the learning subject, and *C* represents the participant unfamiliar with the learning subject. As a result, the proportion of participants *A* and *B* with the highest score in the group is 88.2% (15/17), whereas that of participants *C* with the lowest score in the group is 62.5% (11/17). That is, the more a participant demonstrates a collaborative attitude, the more likely the participant is of being familiar with the learning subject. In addition, we confirmed that there were four sessions (sessions 3, 5, 7, and 9) where the order of the participants' scores was different between the first and last half of the sessions. This indicates that collaborative attitudes are not necessarily consistent through the learning session.

In the following sections, we consider these average scores as the collaborative attitude score of each participant.

4. Analysis of collaborative attitudes based on non-verbal features

4.1. Analysis of non-verbal features

For the non-verbal features that might relate to the participants' collaborative attitudes, previous research[9] used two features: (*a*) the amount of time that a given participant is gazed by other participants, and (*b*) the amount of speech time. In addition, this research considers two other features: (*c*) the amount of gazing time given to other participants, and (*d*) the rate of returned stares from a given participant while being gazed by other participants. These four types of information from all participants (n = 51: three participants in 17 sessions) are extracted by the corpus data (Table 2). Moreover, in order to analyze the difference between participants who are familiar/unfamiliar with the subjects to be learned, six non-verbal features ((*e*) to (*j*) in Table 3) are also analyzed for participants with knowledge of the subject (n = 34: two participants (*A* and *B*) in 17 sessions).

For the analysis, we calculate the correlation coefficient between each participant's score and the value of each feature in Tables 2 and 3. Here, we preliminarily normalized the amount of eye-tracking data because there are differences in the gazing acquisition rate between individuals.

4.2. Analysis results

The right column of Tables 2 and 3 lists the correlation coefficients between the non-verbal features and the collaborative attitude scores. According to these values, feature (*b*) shows a significantly strong correlation ($r > 0.7$, $\alpha < 0.01$). In addition, there are moderate positive correlations between the scores and features (*a*), (*f*), (*h*), and (*j*). Based on the results of Table 2, collaborative participants tend to be noticed by other participants (feature (*a*)), and they speak actively in comparison to the other participants (feature (*b*)). This result provides supportive evidence for the appropriateness of the features used in previous research[9]. On the other hand, because of a weak correlation

Table 2. Non-verbal features and their correlation coefficients (n = 51)

	Feature	Correlation coefficient
(*a*)	Amount of gazed time by other participants	r = 0.377 (α < 0.01)
(*b*)	Amount of speech time	r = 0.721 (α < 0.01)
(*c*)	Amount of gazing time to other participants	r = 0.247 (α < 0.1)
(*d*)	Rate of returning stare while participant is gazed by others	r = 0.179 (*n.s.*)

Table 3. Non-verbal features and these correlation coefficients (n = 34)

	Feature	Correlation coefficient
(*e*)	Amount of gazing time to participant unfamiliar with learning subject (Total)	r = 0.333 (α < 0.1)
(*f*)	Amount of gazing time to participant unfamiliar with learning subject (Speaking)	r = 0.368 (α < 0.05)
(*g*)	Amount of gazing time to participant familiar with learning subject (Total)	r = 0.053 (*n.s.*)
(*h*)	Amount of gazing time to participant familiar with learning subject (Speaking)	r = 0.396 (α < 0.05)
(*i*)	Amount of gazed time by participant unfamiliar with learning subject (Total)	r = 0.310 (α < 0.1)
(*j*)	Amount of gazed time by participant unfamiliar with learning subject (Speaking)	r = 0.480 (α < 0.01)

between the scores and feature (*d*), the action of returning a stare when a given participant is gazed by others is not related to the participants' collaborative attitudes.

Table 3 summarizes the non-verbal features of the participants with knowledge of the learning subject. Based on the results, there is a tendency for the participants to be judged as being more collaborative when they notice the other participants, especially in speaking situations (features (*f*) and (*h*)); moreover, these participants are noticed easily by participants unfamiliar with the subject (feature (*j*)). Meanwhile, there is no correlation between the collaborative attitude scores and feature (*g*).

5. Estimation model of participants' collaborative attitudes

5.1. Estimation model

Based on the results of Section 4, we propose an estimation model of participants' collaborative attitudes. For the modeling parameters, we adopt the features that show significant differences in relation to collaborative attitude scores; these are the features (*a*), (*b*), (*f*), (*h*), and (*j*) shown in Tables 2 and 3. We construct a model that allows an estimate of three levels of collaborative attitude (*high*, *medium*, and *low*) based on multinomial logistic regression analysis. This analysis requires one category of the dependent variable to be the reference category, and it predicts the probability of other possible categories of the dependent variable with the exclusion of the reference category.

We construct two estimation models. One model estimates the three levels of collaborative attitudes based on features (*a*) and (*b*) as independent variables (Model 1). The other model considers the participants' knowledge of the learning subjects (Model 2). This model uses features (*a*), (*b*), (*f*), (*h*), and (*j*) as independent variables for the estimation. The participants are separated into three groups by descending order of their scores (*high*: 16, *medium*: 18, *low*: 17 for Model 1 and *high*: 12, *medium*: 11, *low*: 11 for Model 2).

Table 4. Significance probability (p-value) of each feature: Model 1

	p (high)	*p (low)*
β_a	0.645	0.470
β_b	0.021	0.002

Reference category: *medium*

Table 5. Significance probability (p-value) of each feature: Model 2

	p (high)	*p (low)*
β_a	0.099	0.041
β_b	0.471	0.027
β_f	0.136	0.807
β_h	0.988	0.075
β_j	0.050	0.121

Reference category: *medium*

Equations (1), (2), and (3) represent the results of the multinomial logistic regression model when the category *medium* is set as the reference category. Model 1 is expressed when the regression Equations (4) and (5) are substituted into the variables l and h. In a similar fashion, Model 2 is indicated by substituting Equations (6) and (7) into the variables l and h. These models calculate the probability of three levels of collaborative attitudes by applying the corresponding values into the parameters β_a, β_b, β_f, β_h, and β_j.

$$p_{high} = \frac{e^h}{1+e^l+e^h} \tag{1}$$

$$p_{middle} = \frac{1}{1+e^l+e^h} \tag{2}$$

$$p_{low} = \frac{e^l}{1+e^l+e^h} \tag{3}$$

$$h = 0.002\times\beta_a + 0.037\times\beta_b - 3.816 \tag{4}$$

$$l = -0.006\times\beta_a - 0.059\times\beta_b + 3.768 \tag{5}$$

$$h = -0.029\times\beta_a + 0.030\times\beta_b + 0.163\times\beta_f - 0.001\times\beta_h + 0.138\times\beta_j - 4.920 \tag{6}$$

$$l = -0.051\times\beta_a - 0.086\times\beta_b + 0.025\times\beta_f + 0.116\times\beta_h + 0.118\times\beta_j + 6.404 \tag{7}$$

Table 4 lists the significance probability of each of the features in Model 1. The p-values represent the influence of the differentiation of each feature from the reference category. The amount of speech time (feature (b)) was statistically significant at the 5% level for dividing the *medium-high* level, and 1% for the *medium-low* level. That is, in order not to consider the participants' knowledge level of the learning subject, the results indicate that collaborative attitudes become particularly higher in proportion to the amount of speech time.

Table 6. Result of estimation model on collaborative attitudes

Estimation model	Category	Precision	Recall	F-measure
Model 1	*high*	0.533	0.500	0.516
	medium	0.421	0.444	0.432
	low	0.765	0.765	0.765
	Total	0.571	0.569	0.569
Model 2	*high*	0.667	0.667	0.667
	medium	0.636	0.636	0.636
	low	0.636	0.636	0.636
	Total	0.647	0.647	0.647

Table 5 represents the significance probability of each of the features in Model 2. As a factor that contributes to dividing the *medium-high* level of collaborative attitudes, feature (*j*) is statistically significant at the 5% level. As the factors for the *medium-low* level, features (*a*) and (*b*) are statistically significant at the 5% level. It follows that the more a participant familiar with a subject demonstrates collaborative attitudes in collaborative learning, the more such participant is noted by participants with insufficient knowledge to speak.

5.2. Evaluation results of estimation model

We evaluated the estimation model of the three levels of collaborative attitudes using the machine learning software WEKA[10]. We used 10-fold cross validation by randomly dividing the set of samples into ten parts of approximately equal size. Table 6 lists the estimation results of precision, recall, and f-measure for each category in Models 1 and 2. Here, because the multinomial logistic regression model calculates the probability of each category based on Equations (1) to (7), we deemed each participant's level as the category that exhibits the highest probability.

The results found that the f-measure for Model 1 is 0.569 and for Model 2, where participants are knowledgeable about the learning subject, is 0.647. This means that the estimation accuracy for classifying the participants' collaborative attitudes is approximately two times as correct as for selecting the levels at random (0.33). Moreover, Model 1 shows a high f-measure in the *low* category compared to Model 2; Model 2 correctly estimates all categories. The results indicate that information regarding the participants' knowledge of their learning subjects, in addition to the amount of gaze and speech time, is helpful for estimating the participants' collaborative attitudes.

6. Conclusion

In this paper, we focused on the collaborative attitudes of participants in collaborative learning, and proposed an estimation method by analyzing non-verbal features that might relate to the attitudes in detail. The estimation model detects three levels of the participants' collaborative attitudes based on the multinomial logistic regression analysis. The results of our estimation model showed that the f-measure for classifying the participants' collaborative attitudes is 0.67 when considering the features of the participants' knowledge of their learning subjects. The results of this research support the significance of considering non-verbal information when analyzing collaborative learning interaction. We believe the estimation method contributes to evaluate collaborative learning situations for judging whether each participant truly studies collaboratively or not.

Future directions for this study include the analysis of collaborative learning interaction in greater detail. We intend to annotate verbal information for the interaction corpus (e.g., standard dialog act tags[11] for each utterance). Verbal information allows us to consider the participants' intention, which might enhance the estimation accuracy of the collaborative attitudes. In addition, at the present stage, we only focused on the learning group consisted of three participants, and modeling the collaborative attitudes of each participant. To evaluate whether participants are successfully progressing in their learning, further analysis of group settings and interaction process is required.

Acknowledgements

This work was supported by the Japan Society for the Promotion of Science (JSPS) Grants-in-Aid for Scientific Research (KAKENHI) Grant Nos. 25280076 and 26870588.

References

1. Adelsberger HH, Collis B, Pawlowski JM. *Handbook on Information Technologies for Education and Training*. Springer-Verlag, 2002.
2. Inaba A, Mizoguchi R. Learners' Roles and Predictable Educational Benefits in Collaborative Learning - An Ontological Approach to Support Design and Analysis of CSCL. *In Proc. of 7th International Conference on Intelligent Tutoring System (ITS'04)*, pp. 285–294, 2004.
3. Mori T, Yamada T. Effects of Collaborative Education in the First Year Experience, and its Process: Scaffolding between Students. *Kyoto University researches in higher education*, Vol. 15, pp. 37–46, 2009 (in Japanese).
4. Peña A, De Antonio A. Nonverbal Communication to Support Collaborative Interaction in Collaborative Virtual Environments for Learning. *CEUR Workshop Proceedings*, Vol. 384, 2008.
5. Brennan SE, Chen X, Dickinson CA, Neider MB, Zelinsky GJ. Coordinating cognition: the costs and benefits of shared gaze during collaborative search. *Cognition*, Vol. 106, Issue 3, pp. 1465–1477, 2008.
6. DiMicco JM, Hollenbach K, Bender W. Using visualizations to review a group's interaction dynamics. *In Proc. of CHI '06 Extended Abstracts on Human Factors in Computing Systems*, pp. 706–711, 2006.
7. Kumano S, Otsuka K, Mikami D, Matsuda M, Yamato J. Understanding communicative emotions from collective external observations. *In Proc. of CHI '12 Extended Abstracts on Human Factors in Computing Systems*, pp. 2201–2206, 2012.
8. Hayashi Y, Ogawa Y, Nakano YI. An Experimental Environment for Analyzing Collaborative Learning Interaction. *In Proc. of 15th International Conference on Human-Computer Interaction*, LNCS Vol. 8018, pp. 43–52, 2013.
9. Hayashi Y, Ogawa Y, Nakano YI. Visualization System for Analyzing Collaborative Learning Interaction. *In Proc. of 17th International Conference on Knowledge-Based and Intelligent Information & Engineering Systems (KES2013)*, Procedia Computer Science, Vol. 22, pp. 1379–1387, 2013.
10. Witten IH, Frank E, Hall MA. *Data Mining: Practical Machine Learning Tools and Techniques*. 3rd ed, Morgan Kaufmann, 2011.
11. Jurafsky D, Shriberg L, Biasca D. Switchboard SWBD-DAMSL Shallow Discourse-Function Annotation Coders Manual. *Institute of Cognitive Science Technical Report*, pp. 1–61, 1997.